Yanfei Zhou

Address: 701 W Exposition Blvd, Los Angeles, CA

Email: yanfeizh@usc.edu
Website: yanfeifionazhou.github.io/

Education

- **University of Southern California - Los Angeles, CA** — Aug 2021 - May 2026 (Expected)
  Ph.D. in Data Sciences and Operations (Statistics); Advisor: Prof. Matteo Sesia — *GPA: 4.00/4.00*

- **University of Chicago – Chicago, IL** — Oct 2018 - Jun 2020
  M.S. in Statistics; Advisor: Prof. Weibiao Wu — *GPA: 3.88/4.00*
  Thesis: An empirical study on A Fully Online Approach for Statistical Inference Based on SGD

- **London School of Economics and Political Science - London, UK** — Sep 2015 - Jun 2018
  B.S. with First Class Honours in Statistics with Finance — *GPA: 4.00/4.00*

Research Interests

Uncertainty Quantification; Conformal Prediction; Machine Learning & Deep Learning; Large Language Models

Publications

- **Conformal Classification with Equalized Coverage for Adaptively Selected Groups**
  Yanfei Zhou and Matteo Sesia.
  In Proceedings of NeurIPS 2024 (poster)
  - *Summary*: This paper introduces a conformal inference method to evaluate uncertainty in classification by generating prediction sets with valid coverage conditional on adaptively chosen features. These features are carefully selected to reflect potential model limitations or biases. This can be useful to find a practical compromise between efficiency—by providing informative predictions—and algorithmic fairness—by ensuring equalized coverage for the most sensitive groups. We demonstrate the validity and effectiveness of this method on simulated and real data sets.
  - *Links:* Paper | Code | Poster

- **Conformalized Adaptive Forecasting of Heterogeneous Trajectories**
  Yanfei Zhou, Lars Lindemann, and Matteo Sesia.
  In Proceedings of ICML 2024 (poster)
  - *Summary*: We present a new conformal method for generating simultaneous forecasting bands guaranteed to cover the entire path of a new random trajectory with sufficiently high probability. Prompted by the need for dependable uncertainty estimates in motion planning applications where the behavior of diverse objects may be more or less unpredictable, we blend different techniques from online conformal prediction of single and multiple time series and ideas for addressing heteroscedasticity in regression. This solution is both principled, providing precise finite-sample guarantees, and effective, often leading to more informative predictions than prior methods.
  - *Links:* Paper | Code | Poster

- **Conformal Inference is (almost) Free for Neural Networks Trained with Early Stopping**
  Ziyi Liang, Yanfei Zhou, and Matteo Sesia.
  In Proceedings of ICML 2023 (poster)
  - *Summary*: Although models trained with early stopping empirically often mitigate overfitting and provide more accurate predictions, they generally lack statistical guarantees on their predictive uncertainty unless further calibrated using additional hold-out data. However, naively combining early stopping and conformal calibration is inefficient because it requires multiple splits of the data. To overcome this limitation, we propose conformalized early stopping: a novel method that combines early stopping with conformal calibration while efficiently recycling the same hold-out data. This leads to models that are both accurate and able to provide exact predictive inferences without multiple data splits. We develop practical implementations for different learning tasks, including outlier detection, multi-class classification, and regression.
  - *Links:* Paper | Code | Poster

- **Training Uncertainty-Aware Classifiers with Conformalized Deep Learning.**
  Bat-Sheva Einbinder, Yaniv Romano, Matteo Sesia, and Yanfei Zhou.
  In Proceedings of NeurIPS 2022 (poster)

- *Summary*: We propose a novel training strategy for deep multi-class classifications that yields more informative and reliable uncertainty estimates. The main idea of our method is to mitigate model overconfidence by incorporating information from conformal inference into the training process of deep neural networks. To accomplish this, we introduce a new loss function called the conformal loss. Experiments with synthetic data and real-world data demonstrate that our approach results in smaller prediction sets with higher conditional coverage than existing state-of-the-art alternatives. Additionally, models trained using the conformal loss exhibit greater robustness to overfitting as they achieve higher prediction accuracy on hold-out data.
- *Links:* Paper | Code | Poster

Work in Progress

- **Stance Drift: How AI Middlemen Change What We Mean**
  Lingchong Liu, Yanfei Zhou, Deqing Fu, Jacob Bien, Rachel Wang, Lucy Xia, Xin Tong.
  Preprint (arXiv coming soon)
- **Conformal Inference for Open-Set and Imbalanced Classification**
  Tianmin Xie, Yanfei Zhou, Ziyi Liang, Stefano Favaro, Matteo Sesia.
  Preprint (arXiv coming soon)
- **Conformal Classification with Improved Conditional Coverage via Post Calibration**
  Reza Pinksoo, Yanfei Zhou, Matteo Sesia.
  In preparation
- **Conformal Prediction with Label Differential Privacy**
  Yanfei Zhou, Teresa Bortolotti, Matteo Sesia.
  In preparation

Professional Experience

- **Microsoft - Redmond, WA** May 2025 - Aug 2025
  Data Scientist (Summer internship)
  - Designed and implemented a pioneering Large Language Model (LLM)-powered causal inference framework for decision support, achieving 5–10× efficiency gains vs. baselines at comparable error.
  - Productionized the workflow with automated end-to-end pipelines and built an evaluation harness for sensitivity analyses, ablations, and prompt-stability guardrails.
  - Collaborated with data scientists and the product team to translate model results into business insights reviewed at monthly business reviews with the CVP; authored an internal paper documenting methodology and insights.
- **Amazon AWS – Santa Clara, CA** May 2024 - Aug 2024
  Applied Scientist (Summer internship)
  - Designed and implemented a probabilistic generative system that iteratively produces information-rich questions to evaluate Retrieval-augmented generation (RAG) LLMs, reducing human-in-the-loop effort.
  - Authored an internal paper documenting methodology and preliminary experimental results.
- **CNA Financial Corporation – Chicago, IL** Aug 2020 - Apr 2021
  Data Scientist (Full-time)
  - Built an end-to-end pipeline for fraud detection and changepoint trend analysis of insurance submissions; engineered features and trained regularized generalized linear models to deliver robust estimates under drift.
  - Delivered monitoring notebooks and lightweight dashboards for underwriting; findings informed weekly reviews and guideline updates.
- **CCC Information Services Inc. - Chicago, IL** Jun 2019 - Sep 2019
  Data Scientist (Summer internship)
  - Developed a scalable record-linkage algorithm linking vehicle-damage and casualty data, and introduced an empirical-Bayes estimator to quantify component-level damage uncertainty.

Teaching and Service

- **Instructor.** BUAD 425: Introduction to Business Analytics. Fall 2024
- **Guest Speaker.** DSO 575: Driving Business Transformation with Generative AI and Machine Learning. Fall 2024
- **Teaching Assistant.** BUAD 425: Introduction to Business Analytics. Spring 2024.
- **Teaching Assistant.** GSBA 524: Data Science for Business. Fall 2023.
- **Teaching Assistant.** DSO Summer Scholars and Jumpstart Program. Summer 2023.
- **Conference Reviewer.** NeurIPS 2024–2025, ICML 2025, AISTATS 2025.

Awards

- **University Outstanding TA Award (3 awardees/year).** University of Southern California. 2025.
- **Marshall PhD Fellowship (3 awardees/year).** USC Marshall School of Business. 2025.
- **SEEDS Conference Best Poster Award.** USC Marshall School of Business. 2025.
- **SEEDS Conference Best Poster Award.** USC Marshall School of Business. 2024.
- **Department Master's Program Scholarship.** University of Chicago. 2018-2020.

Skills Summary

- **Programming languages:** Python, PyTorch, R, SQL, Git
- **Platforms/Infra:** AWS, GCP, Azure ML, Hugging Face
- **Methods:** Statistical modeling, causal inference & experiment design, predictive fairness, time-series modeling, hypothesis testing, Bayesian inference, Markov chain Monte Carlo
- **Languages:** English (proficient); Mandarin Chinese (native)